

APPLICATION  
FOR  
UNITED STATES LETTERS PATENT

TITLE: VIRTUAL REALITY PRESENTATION

APPLICANT: RAYMOND C. KURZWEIL

Kenneth F. Kozik  
Fish & Richardson P.C.  
225 Franklin Street  
Boston, MA 02110-2804  
Telephone: 617 542-5070  
Facsimile: 617 542-5906

CERTIFICATE OF MAILING BY EXPRESS MAIL

Express Mail Label No. EL 485 518 372 US

I hereby certify under 37 CFR §1.10 that this correspondence is being deposited with the United States Postal Service as Express Mail Post Office to Addressee with sufficient postage on the date indicated below and is addressed to the Commissioner for Patents, Washington, D.C. 20231.

Date of Deposit 12-21-01

Signature Leroy Jenkins

Typed or Printed Name of Person Signing  
Certificate Leroy Jenkins

## Virtual Reality Presentation

### **CLAIM OF PRIORITY**

This application claims priority under 35 USC §119(e) to U.S. Patent Application Serial No. 60/258,293, filed on December 26, 2000, the entire contents of which are hereby incorporated  
5 by reference.

### **TECHNICAL FIELD**

This invention relates to virtual reality presentations.

### **BACKGROUND**

Virtual reality is the simulation of a real or imagined  
10 environment that can be sensed. One particular type of virtual reality is that experience that can be visually sensed in three dimensions and may additionally provide an interactive experience. A sensory experience in virtual reality may include a visual experience with full real-time motion possibly  
15 accompanied by sound and tactile and other forms of sensory perception.

One form of virtual reality is a three-dimensional (3-D) image that can be explored interactively at a personal computer. With this form of virtual reality, one can control the  
20 experience by manipulating keys or a mouse so that the content of the image moves in some direction or zooms in or out. As the images become larger and interactive controls more complex, the

perception of "reality" increases. More sophisticated efforts involve such approaches as wrap-around display screens and haptics joystick devices that let one feel the display images.

VRML (Virtual Reality Modeling Language) is a language for  
5 describing 3-D image sequences and possible user interactions to go with them. Using VRML, one can build a sequence of visual images into Web settings with which a user can interact by viewing, moving, rotating, and otherwise interacting with an apparently 3-D scene. For example, one can view a room and use  
10 controls to move about the room, as one would experience it if one were walking through it in real space.

#### SUMMARY

In an aspect, the invention features a virtual reality presentation method including capturing motion of a user,  
15 capturing audio of the user, transforming the audio of the user into a different entity, and animating a character with the motion and transformed audio in real-time.

Embodiments may include one or more of the following. The method may also include displaying the animated character on an  
20 output device.

Capturing motion may include attaching multiple motion tracking sensors to areas of the user to track the user's movements, and transmitting signals representing the movements

from the sensors to a computer system. Capturing audio may include attaching a microphone to the user. The microphone may be a wireless microphone.

Transforming the audio may include altering pitch  
5 characteristics of the audio of the user.

Transforming the audio may include transforming the audio into the different entity that is of the opposite gender as the gender of the user.

Animating may include applying the motion to a three  
10 dimensional (3-D) model, and combining the transformed audio to the 3-D model.

In another aspect, the invention features a presentation method including generating a three-dimensional (3-D) model of a character, capturing motion of a user in real-time, capturing  
15 audio of the user in real-time, modifying a gender of the audio of the user, and animating the 3-D model with the motion and modified audio of the user in real-time.

Embodiments may include one or more of the following. The method may also include displaying the animated 3-D model on an  
20 output device.

Capturing motion may include attaching multiple motion tracking sensors to areas of the user to track the user's movements, and transmitting signals representing the movements from the sensors to a computer system.

Capturing audio may include attaching a microphone to the user and the microphone may be a wireless microphone.

Modifying may include altering pitch characteristics of the audio of the user.

5        In another aspect, the invention features a presentation system including a motion tracking device connected to a user, an audio receiving device connected to the user, an audio receiver/converter for receiving and converting the audio into a different gender than that of the user, and a system to produce  
10        an animated three-dimensional character from the motion and converted audio.

15        In another aspect, the invention features a presentation method including detecting motion of a user, detecting audio of the user, altering the audio of the user, synchronizing the motion of the user to an animated character, and synchronizing the altered audio of the user to the animated character.

Embodiments may include one or more of the following.

20        Detecting motion may include receiving signals representing motions from sensors attached to the user and processing the signals in a computer system.

Detecting audio may include receiving audio signals from a microphone attached to the user.

Altering the audio may include a fundamental frequency of the audio.

The method may also include displaying the animated character on an output device. The output device may be a projector, a flat panel plasma monitor, a multi-scan presentation monitor, an electronic whiteboard or a projection  
5 screen.

Embodiments may include one or more of the following. The system may also include an output device.

The motion tracking device may include a set of interconnected sensors affixed to the user, and a transmitting  
10 device for receiving signals from the sensors and sending them to a computer system.

The audio receiving device may be a microphone and the microphone may be a wireless microphone.

The audio receiver/converter may include an audio effects  
15 digital signal processor.

Embodiments of the invention may have one or more of the following advantages.

Real-time voice and motion are combined and displayed as an animated character. A gender of the user is transformed on the  
20 display to the opposite gender with appropriate voice alterations.

Real-time voice is altered and mixed to match a desired output of the user.

The motion of, for example, a performer, is picked up by the sensors attached to the performer and communicated to a computer. The character being generated through the performance animation is animated by the computer and caused in real time to have the same movement as the performer (and as picked up by the sensors). The animated character, which may be a different gender, moves just like the performer. Meanwhile, the voice of the performer is changed, for example, to a different gender through modification of a fundamental frequency.

A mouth of the animated character is caused to move in synchrony with the speaking of the performer.

Other features, objects, and advantages of the invention will be apparent from the description and drawings, and from the claims.

#### DESCRIPTION OF DRAWINGS

FIG. 1 is a block diagram of a virtual reality presentation system.

FIG. 2 is a block diagram of a motion tracking system.

FIG. 3 is a block diagram of a virtual reality presentation process.

FIG. 4 is a block diagram of a graphical rendering process.

FIG. 5 is a block diagram of an audio transformation process.

FIG. 6 is a flow diagram of a virtual reality process.

#### DETAILED DESCRIPTION

Referring to FIG. 1, a virtual reality presentation system 10 includes an input system 12, a computer system 14, and an output system 16. In an example, the input system 12 includes a motion tracking system 18 that is attached to a user 20 that tracks movements of the user 20. The input system 12 also includes an audio voice detector 22 to pick up voice from the user 20.

The computer system 14 includes a memory 24, a processor 26 and a storage system 28. Memory 24 stores an operating system ("OS") 30, a TCP/IP protocol stack 32 for communicating over a network 34 and machine-executable instructions executed by processor 26 to perform a virtual reality (VR) presentation process 36.

The output system 16 may include any suitable presentation display equipment, such as a projector, flat panel plasma monitor, multi-scan presentation monitor, electronic white board or projection screen.

Referring to FIG. 2, the motion tracking system 18 is shown. The motion tracking system 18 may be any commercially available motion tracking system, such as the Wireless MotionStar™ System from Ascension Technology of Burlington, VT.



The motion tracking system 18 includes multiple sensors 50 being placed on the user 20, each of the sensors 50 located on different limbs 52 or other areas 54 of the user's body. A transmitter 56 transmits signals that are received by the sensors 50 and processed by the computer system 14. Computer system 14 may be directly linked to the motion tracking system 18, or connected via a wireless link (not shown).

The audio detector 22 includes a microphone 56 attached near the face of the user 20. In an example, microphone 56 is a wireless head worn microphone, such as the Shure Model WCM16 (from Shure, Inc. of Evanston, IL) head-worn electret condenser microphone designed for wireless use by performers, lecturers, and others who need high quality voice pickup with maximum mobility. The microphone 56 has a wide frequency response, low RF susceptibility and reliable operation at temperature and humidity extremes to make it suitable for any vocal application.

Referring to FIG. 3, the VR presentation process 36 includes a graphical rendering process 70 and an audio transformation process 72. The graphical rendering process 70 generates animated models for real-time animation. The audio transformation process 72 generates naturally sounding vocal harmonies in real-time.

Referring to FIG. 4, the graphical rendering process 70 includes a motion capture subprocess 80 and a three-dimensional

(3-D) modeling subprocess 82. The motion capture subprocess 80 receives real-time data from the motion tracking system 18 attached to the user 20. The 3-D modeling subprocess 82 generates a 3-D model whose movements are capable of being  
 5 driven by user motion data received by the motion capture subprocess 80. The 3-D model is a computer-generated character (animated model) for real-time animation. Any computer-generated character may be used, such as those generated by Kleiser-Walczak of North Adams, MA. The graphical rendering  
 10 process 70 combines in real-time the 3-D character generated by the 3-D modeling process 82 with motions captured in the motion capture subprocess 80 to generate an animated character for display on the output system 16.

Referring to FIG. 5, the audio transformation process of the virtual reality presentation process 36 includes a pitch  
 15 correction subprocess 90, a gender modification subprocess 92 and a mixing subprocess 94. The pitch correction subprocess 90 receives voice data from the microphone 56 and changes the pitch of the input audio signal to a higher or lower level as desired.  
 20 More specifically, the pitch correction subprocess 90 takes an input signal and increases the input signal in frequency by an amount determined by the user 20, typically over a continuously variable range between 0.25 and 4.0 corresponding to shifts of -2 to 2 octaves.

The gender modification subprocess 92 receives the pitch corrected voice data from the pitch correction subprocess 90. The gender modification subprocess 92 produces natural sounding vocal harmonies in real-time in a gender-modified voice. In one  
5 example, the gender modification subprocess 92 converts male voice data into female voice data. In another example, the gender modification subprocess 92 converts female voice data into male voice data. The gender modification subprocess 92 utilizes audio effects digital signal processing, such as found  
10 in Orville™ from Eventide, Inc. of Little Ferry, NJ.

The mixing subprocess 94 receives data generated in the gender modification subprocess 92. The mixing subprocess 94 generates a linear sum of multiple input signals and combines voice and music sources. Any suitable audio mixing algorithm  
15 may be employed, such as from Yamaha Corporation of MI.

Referring again to FIG. 1, the VR presentation process 36 combines audio and visual data and generates an image on output device 16 in real-time. The audio and motions of the user 20 control the displayed image's audio and motions.

20 Referring to FIG. 6, a virtual reality (VR) process 100 includes generating (102) a three-dimensional (3-D) model representation of a character. The character is capable of being animated on a display. The VR process 100 captures (104)

motions of a user having receivers affixed to them graphically and renders (106) the motions to the character.

The VR process 100 receives (108) audio signals from a microphone or other audio pickup device affixed to the user and  
5 modifies (110) a gender of the user. The VR process 100 (112) mixes (112) voice and other audio and outputs (114) an animated character representative of the motions and audio of the user in real-time on the output system 16.

Other embodiments are within the scope of the following  
10 claims.